On Rank-Biased Overlap with Finite and Conjoint Domains

Oscar Kriebel

Supervisor(s): Julian Urbano, Matteo Corsi

EEMCS, Delft University of Technology, The Netherlands

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 23, 2024

Name of the student: Oscar Kriebel
Final project course: CSE3000 Research Project
Thesis committee: Julian Urbano, Matteo Corsi, Mathijs Molenaar

Abstract

Rank-biased Overlap (RBO) is a measure that is used to compare two rankings against each other mathematically using a hyperparameter for persistence, p, to define the importance of items higher up in the rankings. This is able to follow the properties of incompleteness, indefiniteness, and top-heaviness for its results, making it a flexible option for rank similarity. In traditional RBO, the intersection of the items in each of the rankings is weighted by the persistence to reach the final value for RBO as it tends towards infinity. RBO has several assumptions, such as on what happens when rankings are tied, having an infinitely long ranking, and a degree of conjointness between the rankings. In this paper, two new variations are derived on the aspects of having the rankings be fully conjoint, as well as the aspect of having a known finite domain for the rankings. These are described through the equations of RBO^c, for fully conjoint rankings, and RBO^f, for rankings within a known finite domain. While RBO^c tends to be slightly larger and RBO^f tends to be smaller when compared to traditional RBO, both can be more fitting depending on the greater context of their use cases.

Key words: Rank-biased overlap, Ranking similarity, Summation

1 Introduction

In everyday life, rankings are used more often than one would think. Rankings are a list of items in a specific order to show a level of preference or importance. Found in search engines, magazines, and statistics, the ability to compare two different rankings mathematically and know how similar they are is needed more than ever before.

Rank-Biased Overlap [3] is a measure that is used to compare two different rankings. The unique aspect of this measure, over other rank similiarity measures such as Kendall's τ [2] or Average Precision [5], is that it is able to deal with three different properties of rankings. These properties are known as indefiniteness, incompleteness, and top-heaviness.

In short, indefiniteness is the concept that a ranking can be cut off at any point, making their length arbitrary. Next, incompleteness is the principle which states that a ranking may not contain all possible items within its domain. And finally, top-heaviness is the property which states that items higher up on a ranking are more important than lower items.

Together, the properties of indefiniteness and incompleteness lead to the idea of conjointness, or how many items the rankings share altogether. This is due to the possibility that not all items may be ranked by both rankings.

While RBO manages to deal with all three properties, the aim of this paper is to formulate new equations based on RBO to fit specific situations, by setting the goal: '*Define RBO for fully conjoint and/or finite rankings*'. This relates to both the properties of indefiniteness and incompleteness, and the changes that occur when both fully conjoint and finite domains are considered.

This leads to a few distinct options based on the combination of these conditions. These different possibilities are as follows:

1. Conjoint rankings within an infinite domain.
2. Two rankings within a finite domain at a level of conjointness.

In order to properly answer each of these questions, two sub-questions have been defined to help answer these possibilities:

- What occurs in the RBO measure in the case of fully conjoint rankings rather than assuming disjointness?
- How does the RBO measure change when there is a known finite domain and how does it change with both a conjoint and non-conjoint domain?

In this paper, new equations will be defined on how the RBO measure changes for these two possibilities, in the form of RBO^c for the first sub-question and RBO^f for the second sub-question. These changes are then shown through experiments done comparing RBO and the two new variations, using synthetic data to show the difference.

2 Rank-Biased Overlap

When calculating Rank-Biased Overlap [3], there are two different parts that can be looked at: the agreement, A_d at depth d, and their associated weights, w_d.

The agreement A_d can be described as the proportion of items that two rankings, S and L, share at a depth d. We use the term X_d to define the intersection between rankings S and L at depth d. This can be written as:

$$A_{S,L,d} = \frac{|S_{:d} \cap L_{:d}|}{d} = \frac{X_{S,L,d}}{d} \tag{1}$$

Its associated weight w_d is also described using a hyperparameter p for persistence, which can be tuned to fit the amount of top-heaviness wanted. This is described using the infinite sum:

$$\frac{1}{1-p} = \sum_{d=1}^{\infty} p^{d-1} \tag{2}$$

In traditional RBO, we assume that there is an infinite number of items that could appear in each ranking. This leads to a general equation by putting together the agreement, A_d, and its respective weight, w_d, at depth d:

$$RBO(S, L, p) = (1-p) \sum_{d=1}^{\infty} A_{S,L,d}\, p^{d-1} \tag{3}$$

where S is the ranking of shorter length, L is the ranking of longer length with s and l as the lengths of S and L respectively. This is bound within the range of [0,1].

These rankings can be split into two different parts: a seen section and an unseen section. Since RBO is assumed to have an infinite number of items, the unseen section is reached after we have passed the depth equal to the length of the ranking. This also leads to a possible overlap between the seen section of one ranking and the unseen section of another ranking.

d =	1	2	3	4	5	6	7	8	9	10	11	12	...
L =	⟨**a**	**b**	**c**	**d**	**e**	**f**	**g**	**h**⟩	i	j	k	l	...
S =	⟨**b**	**d**	**m**	**n**	**a**	**e**⟩	h	g	c	o	f	i	...

Table 1: Example pair of rankings. Brackets represent the seen items in the rankings and the given S and L where s = 6 and l = 8. For RBO^c, all unmatched items in each ranking would appear at later depths past 12. For RBO^f, we can end at depth d = 12 whereby n = 12 and ϕ would equal 0.75.

When we calculate Rank-Biased Overlap from the seen items in each ranking, we can express the agreement from these rankings based on whether or not the item in one ranking appears in the other. In equation 1, we can see this expressed through the intersection of the two rankings, meaning that this agreement added is either 0 or 1. However, when the unseen section is reached, we cannot know if an unseen item in one ranking appears in the other. This leads to the interpretation that a fractional agreement can be added based on the assumptions made for how this agreement behaves within the unseen section. In traditional RBO, the agreement is assumed to remain the same once it reaches the unseen section.

Eventually, this leads to a final extrapolation of RBO with different length lists. This extrapolation is what we expect the final value to be, which is based on assumptions made on how the agreement behaves. This can be written as:

$$\begin{aligned} RBO_{EXT}(S, L, s, l, p) = & \frac{1-p}{p}[\sum_{d=1}^{l} \frac{X_d}{d} p^d \\ & + \sum_{d=s+1}^{l} \frac{X_s(d-s)}{ds} p^d \\ & + \sum_{d=l+1}^{\infty} (\frac{X_l - Xs}{l} + \frac{X_s}{s}) \frac{p^d}{d}] \end{aligned} \tag{4}$$

This extrapolated value falls between a range, which is calculated from a minimum value, RBO_{MIN}, and a maximum value, RBO_{MAX}. How these values are calculated can be found in the original paper [3] as they serve as a useful reference for the extrapolation.

Altogether, Rank-Biased Overlap makes several assumptions in its calculation, such as having an infinite domain, letting the agreement remain constant after the unseen section is reached, and the non-existence of ties. The first two assumptions are what we are looking to change in this paper.

2.1 Fully Conjoint Infinite Domains

For the first variation covered in this paper, we look at traditional RBO but with a different assumptions for how the agreement behaves within the unseen section of the rankings. Here, we are now assuming that the two rankings are fully conjoint, meaning that as the depth approaches infinity, the agreement will equal 1.

One problem with this variation is the way that the agreement can be modeled to behave in the unseen section. There are multiple ways that the agreement can be modeled, so we would need to choose one that best fits our assumptions. In the end, it should be possible to define a new RBO measure that fits the assumptions related to full conjointness.

2.2 Finite Domains

For the second possibility, the assumption of an infinite domain for the rankings will need to be altered. This leads to several questions that need to be answered to define the equation for RBO in a finite domain, $\mathcal{D}$. This is in regards to the properties of:

- Normalization of weights.
- Degree of conjointness, ϕ.
- Length of the domain.

As part of the journey to define this new RBO variation, each of these problems will need to be addressed in the final equation, with any assumptions being stated along the way.

3 Variations on Rank-Biased Overlap

In order to answer the problem that has been proposed, we need to look at the math behind Rank-Biased Overlap and see where we need to change our assumptions. This will then lead to new equations being made for both variations: assuming full conjointness and assuming a finite domain.

3.1 Fully Conjoint Rank-Biased Overlap

For this variation, we want to change the assumption of how the agreement behaves within the unseen section. In the case of full conjointness, this means that both rankings eventually share every item. Here we will look at how it is defined for the minimum possible value, maximum possible value, and a possible extrapolated value.

3.1.1 Minimum RBO_c

For the minimal value, we can assume that every item in the ranking will appear in reverse order within the unseen section. This essentially means that every item in one ranking will take as long as possible to appear in the other, because every item will need to appear eventually. This is because the two rankings are fully conjoint and share all of the same items.

However, because there is an assumed infinite number of items that are within each domain, then there can always be another pair of unmatched items at each new depth. This means that the agreement will never increase at each depth, only remaining at what we had seen from the seen section of the rankings.

Using this, we can then define the minimum value for the fully conjoint RBO to be equal to RBO_{MIN} as defined in the original paper by Webber et Al. [3]. This minimum value is derived from having the agreement never increase within the unseen section, only ever increasing from the seen items.

3.1.2 Maximum RBO_c

For the maximum value, we can assume that every unmatched item in the ranking will appear as soon as possible in the unseen section of each ranking. This means that by depth $l + s - X_l$ all items that can be matched have been matched as well. Once every item has been matched, then every new item in the unseen section will be matched to itself, leading to the agreement staying at 1 as the depth approaches infinity.

Using this, we can then define the maximum value for the fully conjoint RBO to be equal to RBO_{MAX} as defined in the original paper by Webber et Al. [3].

3.1.3 Extrapolated RBO_c

To figure out what the extrapolation would be based from traditional RBO, we need to change how it handles the agreement from the section of $l+1$ to infinity. From the equation 4, this agreement is described by the final section:

$$A_{l+1:\infty} = \frac{X_l - X_s}{l} + \frac{X_s}{s} \tag{5}$$

From here, we can add on a fraction of the remaining agreement until it equals 1 as the depth tends towards infinity. For this, we will use:

$$lim_{d \to \infty} \frac{d-l}{d} = 1 \tag{6}$$

We can multiply this by the remaining amount of agreement from equation 5 to get a total agreement of 1 when added together as the depth approaches infinity. To use this function, we make the assumption that because the two rankings are fully conjoint, they will share the same items quickly. Other equations can be used to model the behavior of the agreement, but this is used due to its simplicity. From depth d to infinity, the remaining agreement can be described using:

$$(1 - \frac{X_l - X_s}{l} - \frac{X_s}{s})\frac{d-l}{d} \tag{7}$$

This can then be added with the agreement from equation 5 as the total agreement from $l+1$ to infinity. This can then be substituted back into the last section of equation 4. Rearranging this new equation, we get:

$$RBO^c_{EXT} = p^l + \frac{1-p}{p}(\sum_{d=1}^{l} \frac{X_d}{d} p^d + \sum_{d=s+1}^{l} \frac{X_s(d-s)}{ds} p^d$$
$$+ (X_l - X_s + \frac{l(X_s - s)}{s})(ln[\frac{1}{1-p}] - \sum_{d=1}^{l} \frac{p^d}{d})) \tag{8}$$

There are other equations that could be used to model how the agreement behaves in the unseen section, but this is the most straight-forward in its approach as it simplifies more easily. Other models, such as substituting $1 - e^{-pd}$ for equation 6, can fit the same assumptions but with a different behavior based on how the agreement is wanted to be modeled.

3.2 Finite Rank-Biased Overlap

With a finite domain of which the rankings take place in, we can define a new hyperparameter of $n = |\mathcal{D}_S| = |\mathcal{D}_L|$, which is the length of the domain. This does not necessarily mean that the two rankings will share all of the same items, but rather that there is the same number of items that the rankings share in total. This means that $S \in \mathcal{D}_S$ and $L \in \mathcal{D}_L$, thus $n \geq |L|$ must hold as well.

Starting from equation 2 we will need to redefine how the weights are calculated, ensuring that the range of RBO is kept to [0,1]. This gives us:

$$\sum_{d=1}^{n} p^{d-1} = \frac{1-p^n}{1-p} \tag{9}$$

Changing the limit of equation 3 to n and substituting in the the weights from equation 9, we get:

$$RBO^f(S, L, p, n) = \frac{1-p}{1-p^n} \sum_{d=1}^{n} A_d p^{d-1} \tag{10}$$

In the case that $n = |L|$ with even rankings, where $|S| = |L|$, we can just use equation 10 which leads to a version of Average Accuracy [4] with the weight being taken from equation 9.

To find out the extrapolation of this equation with both the seen and unseen sections, we can split it up into three parts: 1) from 1 to s, 2) from $s+1$ to l, and 3) from $l+1$ to n. All of these parts are multiplied by $\frac{1-p}{1-p^n}$ to normalize the final values.

$$\frac{1-p}{1-p^n}[\overbrace{\sum_{d=1}^{s} A_d p^{d-1}}^{1} + \overbrace{\sum_{d=s+1}^{l} A_d p^{d-1}}^{2} + \overbrace{\sum_{d=l+1}^{n} A_d p^{d-1}}^{3}] \tag{11}$$

For each of these sections, RBO^f_{MIN}, RBO^f_{MAX}, and RBO^f_{EXT} will be determined for the minimal, maximal, and extrapolated values from the finite rankings respectively. This can give us a range between $MIN - MAX$ for what the final value could be.

This is used with a new hyperparameter, ϕ, to represent the degree of conjointness between the two domains that make up the rankings. Here ϕ is equal to 0 when the domains are fully disjoint and 1 when the domains are fully conjoint. This can be defined as:

$$\phi_{S,L} = \frac{\mathcal{D}_S \cap \mathcal{D}_L}{n} \tag{12}$$

This is used to create the extrapolation of equation 10, RBO^f_{EXT}. However, it should be the case due to the rankings being drawn from their domains that $1 \geq \phi \geq \frac{X_l}{n}$.

3.2.1 First part: from 1 to s

Using equation 10, we can set it to end at depth s as this will cover the completely seen section from both rankings:

$$\sum_{d=1}^{s} \frac{X_d}{d} p^{d-1} \tag{13}$$

3.2.2 Second part: from $s+1$ to l

This part deals with the remaining section of the seen items in L and the unseen items from S. Here, we have to make assumptions about the agreement added from these unseen items in S to the agreement added from the seen items in L. As explained by Webber et al. [3], this unseen agreement is fractional, rather than being described set-wise through X_l.

For RBO^f_{MIN}:
In this case, all unseen items will be disjoint, so no agreement will need to be added from the unseen items. This leaves us with the agreement from the seen items in L:

$$\sum_{d=s+1}^{l} \frac{X_d}{d} p^{d-1} \tag{14}$$

For** RBO^f_{MAX}**:
Here, all unseen items will be conjoint so they will always match an item from L that we have seen. When a seen item in L matches with an unmatched item in S, it means that there is some item in L that is also unmatched, which the unseen item in S can then match. When the seen item in L does not match with an item in S, then the unseen item in S can then also match it. This gives us the agreement at:

$$\sum_{d=s+1}^{l} \frac{X_d + (d-s)}{d} p^{d-1} \tag{15}$$

Which can then be rearranged to:

$$\frac{p^s - p^l}{1-p} + \sum_{d=s+1}^{l} \frac{X_d}{d} p^{d-1} - s \sum_{d=s+1}^{l} \frac{p^{d-1}}{d} \tag{16}$$

For** RBO^f_{EXT}**:
Here, we want to use the hyperparameter ϕ to determine how the agreement in the unseen section of S behaves. Given that the length of the domain n can equal l, we have to consider how much agreement needs to be added so that the final agreement can equal ϕ whether or not there is a unseen section from L.

To do so, the amount of agreement that needs to be added from $s+1$ to n must equal $\phi - \frac{X_l}{l}$. This amount can be split across the depths from $s+1$ to n equally, added to the seen agreement. This means the contribution from depths $s+1$ to l is equal to:

$$\begin{aligned} &\sum_{d=s+1}^{l} (\frac{X_d}{d} + [\phi - \frac{X_l}{l}] \frac{d-s}{n-s}) p^{d-1} \\ &= \sum_{d=s+1}^{l} \frac{X_d}{d} p^{d-1} + \sum_{d=s+1}^{l} (\phi - \frac{X_l}{l}) \frac{d-s}{n-s} p^{d-1} \end{aligned} \tag{17}$$

When $n = l$, this will give us an agreement of ϕ at depth n, which should be the case due to what we know from equation 12. This can be further rearranged into:

$$\begin{aligned} \sum_{d=s+1}^{l} \frac{X_d}{d} p^{d-1} + \frac{1}{(n-s)(1-p)^2} [p^s(\phi - \frac{X_l}{l}) \\ + p^l[\phi(s-l-1) + \frac{X_l}{l}(l-s+1)] \\ + p^{l+1}(\frac{sX_l}{l} + \phi l - X_l - \phi s)] \end{aligned} \tag{18}$$

3.2.3 Third part: from $l+1$ to n

This part covers the rest of the unseen section for both rankings. Here, we have to make assumptions about how the agreement behaves in the unseen section.

For** RBO^f_{MIN}**:
For this, all unseen items will be disjoint so no agreement will be added. This means we can take the number of shared items from depth l, X_l, and extrapolate that until the final depth:

$$\sum_{d=l+1}^{n} \frac{X_l}{d} p^{d-1} = X_l (\sum_{d=1}^{n} \frac{p^{d-1}}{d} - \sum_{d=1}^{l} \frac{p^{d-1}}{d}) \tag{19}$$

For** RBO^f_{MAX}**:
We can continue the assumption that all unseen items from $s+1$ to l have matched with seen items in L. All remaining items that have not been matched yet will be matched at depth $f = l + s - X_l$. Once that depth has been reached, then the agreement stays at 1 until the depth reaches n:

$$\sum_{d=l+1}^{f} A_d p^{d-1} + \sum_{f+1}^{n} p^{d-1} \tag{20}$$

However, this leads to two different cases: when $f \geq n$ and $f < n$. This is because the length of the domain, n, may not be long enough to allow for all of the unmatched items to appear from $l+1$ to n. Therefore, we must

In the first case, we have to alter equation 22 by removing the second summation and changing the upper limit to n. With some rearranging, we get:

$$\frac{2p^l - 2p^n}{1-p} + (X_l - s - l)(\sum_{d=1}^{n} \frac{p^{d-1}}{d} - \sum_{d=1}^{l} \frac{p^{d-1}}{d}), f \geq n \tag{21}$$

For the second case, the agreement from $l+1$ to f will equal the agreement from equation 15 with an additional 2 for each depth. This lets us redefine equation 22 to:

$$\sum_{d=l+1}^{f} \frac{2d + X_l - l - s}{d} p^{d-1} + \sum_{f+1}^{n} p^{d-1} \tag{22}$$

Which can be rearranged into:

$$\frac{2p^l - p^f - p^n}{1-p} + (X_l - s - l)(\sum_{d=1}^{f} \frac{p^{d-1}}{d} - \sum_{d=1}^{l} \frac{p^{d-1}}{d}), f < n \tag{23}$$

For** RBO^f_{EXT}**:
Here, we can continue the assumption of how the agreement behaves from $s+1$ to l, but now keeping the contribution from the seen section at depth l:

$$\begin{aligned} &\sum_{d=l+1}^{n} (\frac{X_l}{l} + [\phi - \frac{X_l}{l}] \frac{d-s}{n-s}) p^{d-1} \\ &= \sum_{d=l+1}^{n} \frac{X_l}{l} p^{d-1} + \sum_{d=l+1}^{n} [\phi - \frac{X_l}{l}] \frac{d-s}{n-s} p^{d-1} \end{aligned} \tag{24}$$

With some rearrangement, we get:

$$\begin{aligned} \frac{1}{(n-s)(1-p)^2} [p^n(\frac{X_l}{l} + \phi s - \phi n - \phi) \\ + p^l(\frac{nX_l}{l} - X_l - \frac{X_l}{l} + \phi l - \phi s + \phi) \\ + p^{l+1}(\phi s - \phi l + X_l - \frac{nX_l}{l}) \\ + \phi p^{n+1}(n-s)] \end{aligned} \tag{25}$$

3.2.4 Extrapolation in full

Now that we have covered from 1 to n, we can substitute equations 13, 17, and 24 into the general equation 11 for the full extrapolation:

$$RBO^{f}_{EXT}(S, L, s, l, p, n, \phi) = \frac{1-p}{1-p^n}\Big[\sum_{d=1}^{l}\frac{X_d}{d}p^{d-1} + \frac{X_l}{l}\sum_{d=l+1}^{n}p^{d-1} + \sum_{d=s+1}^{n}(\phi - \frac{X_l}{l})\frac{d-s}{n-s}p^{d-1}\Big] \quad (26)$$

With some rearrangement for the final equation, we then get:

$$RBO^{f}_{EXT} = \frac{1-p}{1-p^n}\Big(\sum_{d=1}^{l}\frac{X_d}{d}p^{d-1}\Big) + \frac{1}{(1-p^n)(1-p)(n-s)}\Big[p^s(\phi - \frac{X_l}{l}) + \frac{X_l}{l}p^l(n-s) + \frac{X_l}{l}p^{l+1}(s-n) + p^n(\phi s - \phi n - \phi + \frac{X_l}{l}) + \phi p^{n+1}(n-s)\Big] \quad (27)$$

4 Experiment and Discussion

In order to see how the newly defined RBO^c and RBO^f are different from traditional RBO, we can use synthetic rankings for comparison. 10000 different pairs of rankings are generated from a set of 1000 items for comparison, setting 0.9 for p and 1000 for n from the number of items, using a similar process to Corsi, M. and Urbano, J. [1] while ensuring that no ties are generated. When a value for ϕ is needed for RBO^f, the value of 1 is used as all items from the rankings are pulled from the same 1000 items. Given the definition of ϕ, when the final depth is reached, all items should have appeared in both rankings with full conjointness.

Using the rankings generated, two different graphs have been created, as seen in Figure 1 that compare the value from RBO_{EXT} from equation 4 against RBO^c_EXT and RBO^f_{EXT} from equations 8 and 27 respectively. There is also Table 2 which shows a statistical summary using the absolute difference between the generated values between RBO against RBO^c and RBO^f. The medium differences show the proportion of rankings that have an absolute difference between 0.01 and 0.1 when compared to the RBO measures, while the large difference shows the proportion of rankings with an absolute difference between 0.1 and 1.

Table 2: Summary of the differences between traditional RBO and the variations RBO^c and RBO^f using synthetic data. M for medium differences (0.01,0.1], and L for large in (0.1,1].

	$\lvert RBO - RBO^c\rvert$				$\lvert RBO - RBO^f\rvert$			
p	Avg.	Max.	M	L	Avg.	Max.	M	L
0.8	$3.8\text{x}10^{-3}$	0.02	<1%	<1%	$1.5\text{x}10^{-3}$	0.03	5%	<1%
0.9	$6.4\text{x}10^{-4}$	0.12	2%	<1%	0.01	0.14	29%	1%
0.95	$4.8\text{x}10^{-3}$	0.26	11%	<1%	0.05	0.33	52%	15%

Also, the graph in Figure 2 compares the average RBO value across p in the range of [0.5,0.95]. This helps to show the different impact that p can have on the final RBO value as RBO^f uses a different weight as shown in equation 9.

From these graphs, we can see that the extrapolated values from RBO^c are quite similar to the traditional RBO value, if not for a few values being higher than normal. This makes sense as for RBO^c we are assuming that both rankings are fully conjoint. Even with the weights decreasing as the depth increases, with the agreement only ever increasing more and more as it approaches infinity, this means that there is always be a higher value than compared to the traditional RBO value.

And from what we can see from the values from Table 2, the difference between RBO^c and RBO is quite low, only increasing slightly as p increases. The difference would only be more apparent when the rankings themselves are not very conjoint to begin with, in the case of the seen section having a low agreement. This would then remain the same in traditional RBO while increasing all the way to 1 in RBO^c. This difference could also be attributed to the way that the agreement's behavior is modeled in this paper, with other versions having slightly more or less agreement added at each depth.

For RBO^f, we can see a much larger amount of difference between the traditional and new value, whereby the RBO^f value is either the same or lower than the traditional one. The reason for this could be interpreted as either through the new definition of the weights with the finite summation of p. Given that equation 27 uses a finite domain rather than infinite one, there would not be the section of n to infinity that is added onto the final total. This would help to explain why the

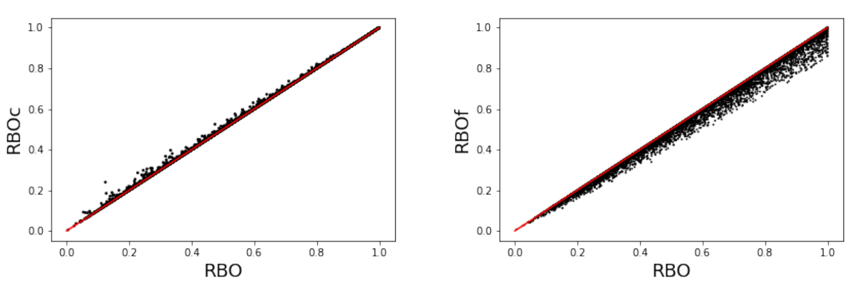

Figure 1: The difference between traditional RBO and the variations RBO^c and RBO^f.

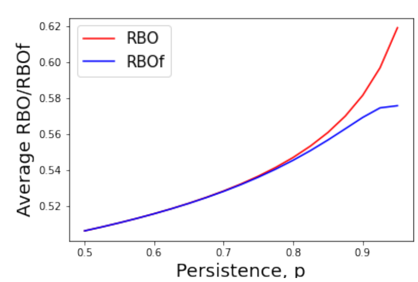

Figure 2: The average difference between RBO and RBO^f across values for p

difference at the large values is greater as a larger value means a higher agreement, one that would normally continue for an infinite depth rather than a finite one. This trend can also be seen in Figure 2, as when the value for p gets larger, the average difference between the two RBO versions gets larger as well, with the one for RBO^f being smaller. Similarly to RBO^c, this can also come from the assumptions related to the behavior of the agreement within the unseen section, where other ways to model this behavior may fit better.

From Table 2, we can also see that the absolute difference increases as well as the persistence increases. Given the definition for the persistence, with a larger p, there is less of a weight put onto the top items. In traditional RBO, this would mean that the later depths have a greater impact on the final value, depths that RBO^f does not reach. This would make it so that the value for RBO^f is more sensitive to what p is selected for use.

5 Responsible Research

In this paper, synthetic rankings were used for comparing the new formulas introduced in this paper to the traditional RBO measure. For these rankings, a seed of 65074241 was used when generating the rankings, meaning that the rankings can be fully reproduced if wanted.

The synthetic rankings that have been generated were made from the code provided from Corsi, M. and Urbano, J[1]. This allows for easy parameterization of the rankings generated so that we can get exactly what we need from the synthetic rankings. The aspect of ties is also ignored when generating the rankings for simplicity.

Given that there is no real data used here, all of the research done in this paper can be reproduced in full using the provided equations in this paper and in the tradiational RBO paper [3].

When this results from this paper are used, all assumptions made in this paper, in regards to the existence of ties, how the agreement is extrapolated, and the properties of the rankings themselves, need to be considered. This paper has specific cases where each of the equations should be used, where if this is deviated from, it can lead to incorrect results. This means that these equations cannot be used in a general scenario, but rather ones where it meets the specific criteria for each of the equations here.

For RBO^c, it should be used when the two rankings come from the same domain. This also uses the assumption that the rankings will have their unmatched items appear in the other ranking quickly, rather than being dispersed across the infinite depth. Then for RBO^f, it should be used when the two rankings come from domains of which all of the items would be known. The domains for these two rankings should also be the same length, but do not need to share all the same items. Here, the unmatched items are assumed to be appear evenly throughout the unseen section as well.

6 Conclusions and Future Work

In this paper, the goal was to '*Define RBO for fully conjoint and/or finite rankings*'. To be able to do this, two sub-questions were defined which led into two different variations for RBO: RBO^c where RBO is assumed to be fully conjoint as its agreement eventually reaches 1, and RBO^f where RBO is assumed to not have an infinite domain and introduces a new hyperparameter, n, for the domain length, and ϕ, for the conjointness of the two rankings' domains. These new assumptions are followed along from the original derivation of RBO which leads to new equations being defined, equations 8 and 27 for RBO^c and RBO^f respectively. The range for the extrapolated values of RBO^c_{EXT} and RBO^f_{EXT} are also provided through their MIN and MAX variations.

These new formulas for Rank-Biased Overlap are then compared to the traditional RBO to see how differently they behave with their final results. From what we see, RBO^c is still quite close to the traditional RBO while still remaining slightly higher due to the agreement being potentially higher, and RBO^f is lower than the traditional RBO which comes from the assumptions related to the behavior of the agreement and the sensitivity of the hyperparameter p.

There is further work that can be done on this topic. For both variations, only one of many different ways to model the agreement within the unseen section is chosen, with other models being perhaps better fits depending on the circumstance. Furthermore for RBO^f, the case of having two different domain lengths has not been covered which may also further impact the final result. Similarly, the values from RBO^f_{MIN} and RBO^f_{MAX} are not tight bounds as phi could be used to determine how many items would appear in the rankings, allowing for a better definition of the range within the known domains.

And for both of these cases, other assumptions that traditional RBO need to also be considered, such as what occurs when there are ties in the rankings. Finally, with the approach that RBO^f has towards the conjointness of the two ranking's domains with the hyperparameter ϕ, this could also be expanded to the traditional RBO to better defined how the extrapolation of the RBO score is made based on further knowledge.

[1] https://github.com/julian-urbano/sigir2024-rbo

References

[1] M. Corsi and J. Urbano. The treatment of ties in rank-biased overlap. *Proc. 47th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2024.

[2] M. G. Kendall. A new measure of rank correlation. *Biometrika*, 1938.

[3] W. Webber, A. Moffat, and J. Zobel. A similarity measure for indefinite rankings. *ACM Transaction on Information Systems*, 2010.

[4] S. Wu and F. Crestani. Methods for ranking information retrieval systems without relevance judgments. *Proc. 18th Annual ACM Symposium on Applied Computing*, 2003.

[5] E. Yilmaz, J. A. Alsam, and S. Robertson. A new rank correlation coefficient for information retrieval. *Proc. 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2008.